Appendix C

Available online at www.sciencedirect.com
SCIENCE DIRECT
ELSEVIER
Metabolic Engineering 6 (2004) 101–108
METABOLIC ENGINEERING
http://www.elsevier.com/locate/ymben

Elevation of seed α-tocopherol levels using plant-based transcription factors targeted to an endogenous locus

Alison L. Van Eenennaam,[a,1,2] Guofu Li,[b,2,3] Mylavarapu Venkatramesh,[c] Charlene Levering,[a] Xiaosong Gong,[b] Andrew C. Jamieson,[b] Edward J. Rebar,[b] Christine K. Shewmaker,[a,4] and Casey C. Case[b,*]

[a] *Monsanto Company, 1920 5th Street, Davis, CA 95616, USA*
[b] *Sangamo BioSciences Inc., Point Richmond Technology Center, 501 Canal Boulvard Suite A100, Richmond, CA 94804, USA*
[c] *Renessen LLC, 3000 Lakeside Drive, Suite 300, Bannockburn, IL 60015, USA*

Received 27 June 2003; accepted 3 November 2003

Abstract

Synthetic zinc finger transcription factors (ZFP-TFs) were designed to upregulate the expression of the endogenous *Arabidopsis* γ-tocopherol methyltransferase (GMT) gene. This gene encodes the enzyme responsible for the conversion of γ-tocopherol to α-tocopherol, the tocopherol species with the highest vitamin E activity. Five three-finger zinc finger protein (ZFP) DNA binding domains were constructed and proven to bind tightly to 9 bp DNA sequences located in either the promoter or coding region of the GMT gene. When these ZFPs were fused to a nuclear localization signal and the maize C1 activation domain, four of the five resulting ZFP-TFs were able to upregulate the expression of the GMT gene in leaf protoplast transient assays. Seed-specific expression of these ZFP-TFs in transgenic *Arabidopsis* produced several lines with a heritable elevation in seed α-tocopherol. These results demonstrate that engineered ZFP-TFs comprised of plant-derived elements are capable of modulating the expression of endogenous genes in plants.

1. Introduction

The tools of plant biotechnology allow the engineering of novel traits via introduction of foreign genes into plants and expression of these genes in a developmental or tissue specific manner using selected promoters. Desired traits can alternatively be obtained by regulating the expression of endogenous genes. One method to achieve this type of targeted gene regulation is through the use of engineered transcription factors. Transcription factors are trans-acting proteins that bind to specific cis-elements and regulate gene expression. Transcription factors are typically modular, consisting of a DNA-binding domain (DBD) and an effector domain (ED) that interacts with other regulatory proteins to either activate or repress transcription. The Cys_2–His_2 zinc finger proteins (ZFPs) are the most common DBDs in eukaryotes and over the past decade this motif has emerged as amenable to manipulations designed to achieve the specific recognition of a predetermined DNA sequence (Pabo et al., 2001, Segal and Barbas, 2001, Beerli and Barbas, 2002). These designer ZFPs have been fused to different EDs to create hybrid zinc finger protein transcription factors (ZFP-TFs) that have been used successfully in the regulation of endogenous chromosomal genes in both animal cell lines (Beerli et al., 2000; Zhang et al., 2000; Bartsevich and Juliano, 2000; Dreier et al., 2001 ; Liu et al., 2001, Ren et al., 2002) and more recently transgenic plants (Guan et al., 2002).

This study reports the production of ZFP-TFs comprised of plant-derived DNA sequences and engineered to activate the endogenous *Arabidopsis*

*Corresponding author.
E-mail address: ccase@sangamo.com (C.C. Case).
[1] Current address: Department of Animal Science, One Shields Avenue, University of California, Davis, CA 95616, USA.
[2] Contributed equally to the paper.
[3] Current address: Agronomic Traits, Trait and Technology Development, Pioneer Hi-Bred International, Johnston, IA 50131, USA.
[4] Current address: BluGoose Consulting, Woodland, CA 95776, USA.

1096-7176/$ - see front matter © 2003 Elsevier Inc. All rights reserved.
doi:10.1016/j.ymben.2003.11.001

γ-tocopherol methyltransferase (GMT) gene (GenBank Accession AF104220). GMT encodes the enzyme responsible for addition of a methyl group to ring carbon 5 of γ-tocopherol to form α-tocopherol (Fig. 1), the tocopherol isoform with the highest vitamin E activity (Bramley et al., 2000). GMT is frequently limiting in seed tissue and as a result the tocopherol composition of seed from many plant species is made up predominantly of γ-tocopherol (Sheppard et al., 1993). Transgenic overexpression of an *Arabidopsis* GMT cDNA was previously found to result in a large increase in α-tocopherol relative to control (Shintani and Della-Penna, 1998).

Five three-finger ZFPs were designed to bind to target 9 bp sequences found in the promoter or coding region of the endogenous *Arabidopsis* GMT gene. These ZFPs were each fused to the maize opaque-2 nuclear localization signal (GenBank Accession M29411) and the maize C1 (GenBank Accession TVZMMB) activation ED (Guyer et al., 1998, Goff et al., 1991) to make ZFP-TFs. Expression of these ZFP-TFs in transgenic *Arabidopsis* under the control of an embryo specific promoter (Kridl et al., 1991) resulted in several lines that had an elevated seed α-tocopherol percentage, with the best line demonstrating a heritable 20 fold increase in percent α-tocopherol relative to control seed. Transgenically modulating endogenous gene expression to achieve a desired phenotype using ZFP-TFs offers a powerful tool to plant biotechnologists (Guan et al., 2002, Ordiz et al., 2002). This is first time that ZFP-TFs derived wholly from plant sequences have been described and found to modulate expression of an endogenous target gene in a whole plant.

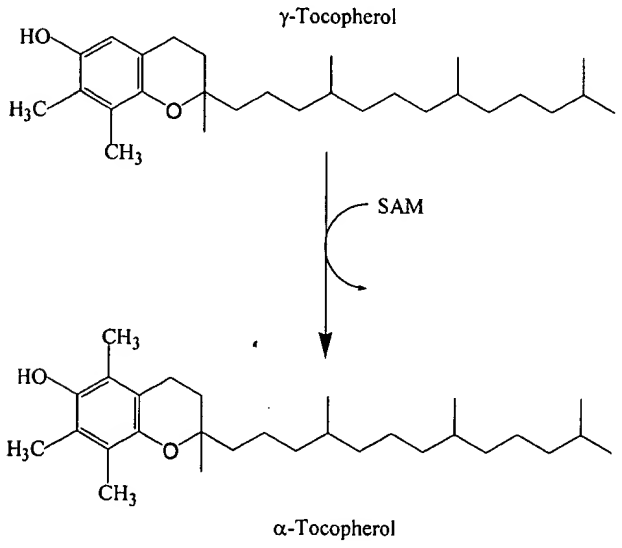


Fig. 1. Biosynthetic reaction catalyzed by GMT. GMT adds a methyl group to ring carbon 5 of γ-tocopherol.

2. Materials and methods

2.1. Nuclei isolation and DNase I hypersensitive mapping

Leaves from 3–4 week *Arabidopsis thaliana* (ecotype Columbia) grown on soil under a 16-h-light/8-h-dark cycle and 25°C constant temperature were harvested and intact nuclei were digested with increasing concentrations of DNase I (Li et al., 1998). Hypersensitivity mapping was performed essentially as described (Zhang et al., 2000). Briefly, DNA from DNase I-treated nuclei was isolated and digested to completion with Dra II, resolved on agarose gels and transferred to nylon membranes. These membranes were hybridized to an indirect end-labeled 500 bp probe derived from the genomic sequence located 1.5–2 kb upstream of the GMT transcription start site.

2.2. Synthesis, purification, and gel shift analysis of plant-derived ZFPs

The strategy used to design, synthesize and purify the ZFP DNA binding domains was as described in Jamieson and Li (2002) and Rebar and Jamieson (2002). The quantitative electrophoretic gel mobility shift assay was performed as described previously (Liu et al., 2001) except that an additional 150 mM NaCl was added to the binding buffer.

2.3. Construction of ZFP-TFs for protoplast based transient assay

ZFP binding domains were subcloned into a plant ZFP expression vector, YCF4, generated from pcDNA3.1 (Invitrogen, Carlsbad, CA). YCF4 contains a CaMV35S promoter driving expression of the coding sequences from the maize opaque-2 nuclear localization signal (RKRKESNRESARRSRRSRYRKKV) and 60 amino acids from the maize C1 activation domain (AGSSDDCSSAASVSLRVGSHDEPCFSGDGDGD WMDDVRALASFLESDEDWLRCQTAGQLA). All of the ZFP-TF expression vectors were constructed by subcloning the ZFP fragments into the KpnI and BamHI sites in YCF4 between the NLS and the C1-ED (Fig. 4A).

2.4. Isolation and transformation of Arabidopsis leaf protoplasts

Isolation and transformation of *Arabidopsis* leaf protoplasts were carried out as described (Abel and Theologis, 1994) with the following modifications. Purified protoplasts were resuspended to a density of 5×10^6 protoplast/ml in a solution containing 400 mM mannitol, 15 mM $MgCl_2$ and 5 mM Mes-KOH at pH 5.6. Approximately 1.6×10^6 protoplasts (300 μl

suspension) were added to a mixture of 50 μg ZFP-TFs plasmid DNA in 15 ml disposable conical tube and mixed well. PEG-CMS solution (400 mM mannitol, 100 mM $Ca(NO_3)_2$, 40% PEG3350) was immediately added to a final PEG concentration of 20% and the suspension was carefully mixed to a homogeneous phase. Following incubation at room temperature for 30 min, the transfected mixture was diluted by 10 ml of protoplast growth medium (400 mM sucrose, 4.4 g/L Murashige and Skoog salt and vitamin mixture (Gibco, Rockville, MD), and 250 mg/L xylose) and transferred to a 10 cm petri dish and incubated in the dark at 25°C for 18–24 h.

2.5. Analysis of endogenous GMT activation: TaqMan analysis

For TaqMan quantitative RT-PCR analysis of mRNA abundance in protoplasts, total RNA from transfected *Arabidopsis* leaf protoplast was isolated using the plant RNeasy kit (Qiagen, Valencia, CA). Real-time PCR analysis was performed in a 96-well format on an ABI 7700 SDS machine (Perkin–Elmer, Chicago, IL) and analyzed with SDS version 1.6.3 software. RNA samples (5 ng) were mixed with 0.3 μM each primer, 0.1 μM probe, 5.5 mM $MgCl_2$ and 0.3 mM each dNTP, 0.625 unit of AmpliTaq Gold RNA polymerase (Hoffman La-Roche, Inc.), 6.25 units of Multiscribe Reverse Transcriptase, and 5 units of RNase Inhibitor in TaqMan buffer A (Perkin–Elmer). The reverse transcription was performed at 48°C for 30 min. After denaturing at 95°C for 10 min, PCR amplification reactions were conducted for 40 cycles at 95°C for 15 s and at 60°C for 1 min. The *Arabidopsis* GMT primer and probe set (AATGATCTCGCGGCTGCT, GAA TGGCTGATTCCAACGCAT, FAM-TCACTCGCT CATAAGGCTTCCTTCCAAGT-TAMRA) were used to measure the *Arabidopsis* GMT expression levels. The GAPDH primer and probe set (GATCATCAAG ATTGTATCTGATC, CGGTTCCTTCGATAACTA AGT, FAM- CAATGCCTAGTTCCCCCCAGGGGAG-TAMRA) were used to monitor the internal control GAPDH mRNA.

For TaqMan analysis of mRNA levels in developing *Arabidopsis* seed, total RNA was isolated from four developing siliques per plant. Total RNA was prepared using the SV Total RNA Isolation Kit (Promega, Madison, WI). RT-PCR was performed as described above, except that the assay consisted of 30 ng total RNA, 0.8 μM each primer, 0.15 μM probe, 10 units of Multiscribe Reverse Transcriptase, and 5 units of RNAse Inhibitor in 1X TaqMan Universal PCR Master Mix w/o AmpErase® (Perkin–Elmer). The primer and probe sets (AATGATCTCGCGGCTGCT, GAATGG CTGATCCAACGCAT, FAM-TCACTCGCTCATAA GGCTTCCTTCCAAGT-TAMRA); (TGCCAGAAC AAGAAGGGTGG, ATACCGACGCCGCCG, FAM-TCGTCCGACGACCCTGCGG-TAMRA); (CGTCC CTGCCCTTTGTACAC, CGAACACTTCACCGGA TCATT, VIC-CCGCCCGTCGCTCCTACCGAT-TA MRA); and (TGCCCCACCTTGAGACAAG, CTTG CTCTGGTTGGTGTTGCT, VIC-CCCTGGAATCT AACGGCCTTGGCA-TAMRA) were used to assay GMT, C1-ED, 18S RNA, and endogenous napin RNA, respectively. For each RNA sample, two replicates were amplified in a one-step reaction and cycle threshold values were obtained. Relative values were calculated using the comparative Ct method described in the manufacturer's User Bulletin (The Perkin–Elmer Corporation, 1997).

2.6. Plant vector construction, Arabidopsis transformation

ZFP-TFs (A-E) were subcloned downstream of the napin embryo-specific promoter (Kridl, et al., 1991) using the *Not*I and *Hin*dIII sites (Fig. 4A) in a plant transformation binary vector. These vectors and their parent (lacking a ZFP-TFs insert as a control) were electroporated into *Agrobacterium tumefaciens* strain ABI and grown under standard conditions (McBride et al., 1994), reconfirmed by restriction analysis, and transformed into *Arabidopsis* using the dipping method (Clough and Bent, 1998). Transgenic T1 plants (12–22 independent insertion events for each ZFP-TF) were grown to maturity and T_2 seed was analyzed for tocopherol content and composition using normal phase HPLC methods as described by Savidge et al. (2002). The four events with the highest seed α-tocopherol percentage from each of two ZFP-TFs (A and B) were advanced to the next generation and T3 seed from 10 individual T2 plants per event was analyzed for tocopherol content and composition.

3. Results and discussion

3.1. DNase I hypersensitive mapping of the Arabidopsis GMT gene

Previous studies indicated that chromatin organization is a determinant of ZFP-TF function within endogenous chromosome loci. The positioning of nucleosomes at endogenous loci can prevent the access of DBDs to the DNA binding site (Li et al., 2001). Targeting ZFP-TFs to an open and accessible region can greatly increase their chances of successfully regulating the target gene (Zhang et al., 2000, Liu et al., 2001). DNase I hypersensitive mapping was performed to locate the regions of the GMT gene accessible to ZFPs. Given the impracticality of obtaining large quantities of *Arabidopsis* embryo tissue, leaf tissue was used for the

DNase I hypersensitive mapping of the GMT gene. These data were then utilized to infer potential accessible sites in the embryo. A previous study demonstrated that some DNase I hypersensitive sites, especially those near the transcription start site, are conserved between different tissues and cell types (Liu et al., 2001).

The DNase I hypersensitive mapping revealed four hypersensitive sites (Fig. 2). The most accessible and therefore obvious hypersensitive site was a doublet centered at −140 bp upstream of the transcription start site. It covered the transcription start site and the entire 5′ UTR. The second hypersensitive site was centered around −600 bp. This hypersensitive site was fairly narrow but had a high intensity. The two other hypersensitive sites (−1000 bp, +500 bp) did not appear to be as accessible to the DNAase I and hence were fainter and more difficult to discern.

3.2. *Biochemical characterization of ZFPs and design of ZFP-TFs*

Five 9 bp potential binding sites were selected based on their positions relative to the hypersensitive sites and their suitability as ZFP target sequences (Fig. 3A). Two of the binding sites were located in the most accessible

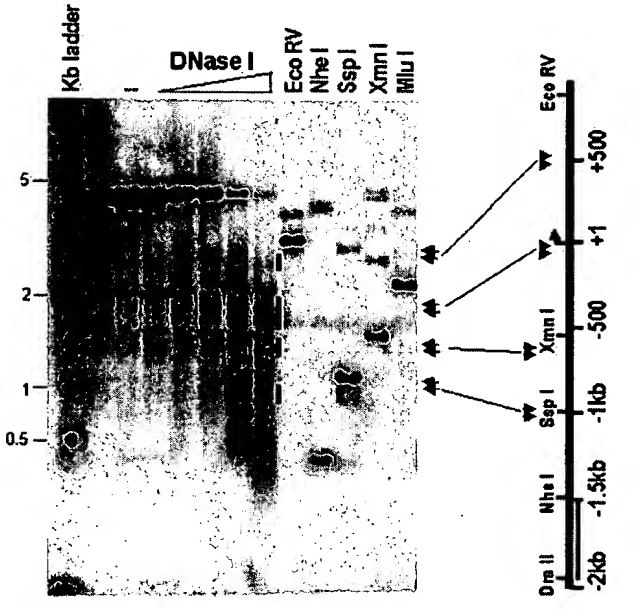
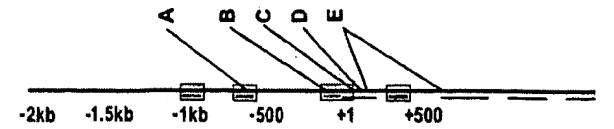


Fig. 2. DNase I digestion profile of the *Arabidopsis* GMT gene. Nuclei from *Arabidopsis* leaves were digested with increasing concentrations of DNaseI or the indicated restriction enzymes. A 500 bp sequence of the genomic sequence located 1.5–2 kb upstream of the GMT transcription start site was used as a probe on a Southern blot of extracted nuclei DNA. Transcription start site (+1) of the GMT gene (vertical graphic) and the location of the probe (black line) are shown. Double arrows indicate the relationship of the observed hypersites to the location of the region on the GMT gene.



(A)

ZFP	Target (5′ to 3′) Sequence	Subsites	Finger designs (-1 to +6)	Gel Shift	Kd (nM)
A	GAGGAAGGGg	GGGg GAAg GAGg	RSDHLAR QSGNLAR RSDNLTR	Bound Free	0.001
B	GGGGAAAGGg	AGGg GAAa GGGg	RSDHLTQ QSGNLAR RSDHLSR	Bound Free	0.0036
C	GAAGAGGGTg	GGTg GAGg GAAg	QSSHLAR RSDNLAR QSGNLAR	Bound Free	0.0036
D	GAGGAGGATg	GATg GAGg GAGg	QSSNLQR RSDNLAR RSDNLQR	Bound Free	0.0003
E	GAGGAGGAGg	GAGg GAGg GAGg	RSDNLAR RSDNLAR RSDNLTR	Bound Free	0.0008
SP1	GGGGCGGGGg	GGGg GCGg GGGg	KTSHLRA RSDELQR RSDHLSK	Bound Free	0.055

(B)

Fig. 3. The location and DNA binding properties of ZFPs targeted to the *Arabidopsis* GMT. (A) The location of the target sites of five ZFP DNA binding domains (A–E) designed to bind to 9 bp sequences in the *Arabidopsis* GMT gene. Numbering is relative to the start site of transcription (+1). Hypersensitive sites inferred from Fig. 2 (gray rectangles) and the first two exons of the GMT gene (black horizontal lines) are indicated. (B) The 9 bp target sequence, amino acid sequence of positions "−1" through "+6" of the α-helix of each zinc finger, and gel shift assays showing the apparent K_d of plant-derived zinc finger protein DNA binding domains (A–E) and SP1 (positive control) for their target sequence.

DNase I hypersensitive sites (A, B), one (C) was at the margin of a hypersensitive site, and the remainder (D and E) were not located in a hypersensitive site. We intentionally evaluated ZFPs designed to bind to sequences that were both within and external to the mapped hypersensitive sites to determine the importance of chromatin infrastructure on in planta activity of the transcription factors.

Although these five ZFPs were assembled using previous described methods (Zhang et al., 2000, Liu et al., 2001) they were distinct from any previously reported synthetic ZFPs in two ways. Firstly, the backbone of these ZFPs was of plant origin having been assembled of DNA fragments derived from several different plant ZFPs (Jamieson and Li, 2002). This unique zinc finger backbone had overall sequence similarities to human SP1, a natural ZFP (Cook et al., 1999). Secondly, while the first two fingers of these ZFPs were canonical C_2H_2 fingers, in the third finger the second histidine and the two adjacent upstream amino acid residues were substituted with GlyGlyCys, making

the third finger a C_2HC type (Rebar and Jamieson, 2002). Remarkably these novel ZFPs with a plant backbone and one C_2HC finger bound to their naked DNA target sequences with higher affinities and specificities than those previously reported (Zhang et al., 2000, Liu et al., 2001) for synthetic ZFPs using a more conventional backbone (data not shown). The amino acids at position "−1" through "+6" of the α-helix of each finger for each ZFP used are shown in Fig. 3B. Gel shift analysis showed that the K_d value of these ZFPs for their naked DNA target sequence ranged from 0.0001 to 0.02 nM. As a comparison, SP1 exhibited a K_d of 0.055 nM for its target under the same gel shift conditions.

Translational fusions were made between the opaque-2 nuclear localization signal (NLS), the five ZFPs, and the activation domain of C1 to generate the plant-derived ZFP-TFs (Fig. 4A). Previous studies of C1, a transcriptional activator of genes encoding biosynthetic enzyme of the maize anthocyanin pigment pathway, demonstrated that the 100 amino acids at the carboxyl-terminal were able to function as a transcriptional activator in maize, yeast and *Arabidopsis* (Guyer et al., 1998; Goff et al., 1991). A shortened C1 activation domain (60 carboxyl-terminal amino acids) was used as the ED in this study as it was found to give superior reporter gene activation in *Arabidopsis* leaf protoplast-based transient assays (data not shown).

3.3. Transcriptional activation of the endogenous Arabidopsis GMT in leaf protoplasts

Plasmids containing each of the five ZFP-TFs (A-E) cloned under the control of the constitutive 35S promoter (Fig. 4A) were transiently transfected into *Arabidopsis* leaf protoplasts and the effect on GMT mRNA levels measured by quantitative real-time PCR analysis. Four (A, B, C and E) of the 5 ZFP-TFs tested increased GMT gene expression greater than two-fold, with B, the most effective activator, giving a 5-fold increase (Fig. 4B). A dosage response experiment for B verified that the level of activation was positively correlated to the amount of DNA used (data not shown). Consistent with earlier studies showing that DNase I hypersensitivity is an indication of accessibility (Utley et al., 1997, Liu et al., 2001), the two ZFP-TFs (A and B) targeted to DNase I hypersensitive sites were found to be effective activators. Interestingly, one of the two ZFP-TFs (D and E) targeted to sites outside of the DNase I hypersensitive sites were also effective activators of the GMT gene in leaf protoplasts. These results are not entirely unforeseen as it has been previously shown that DNase I hypersensitive mapping is not sufficiently precise to reveal small stretches of accessible DNA (Zhang et al., 2000).

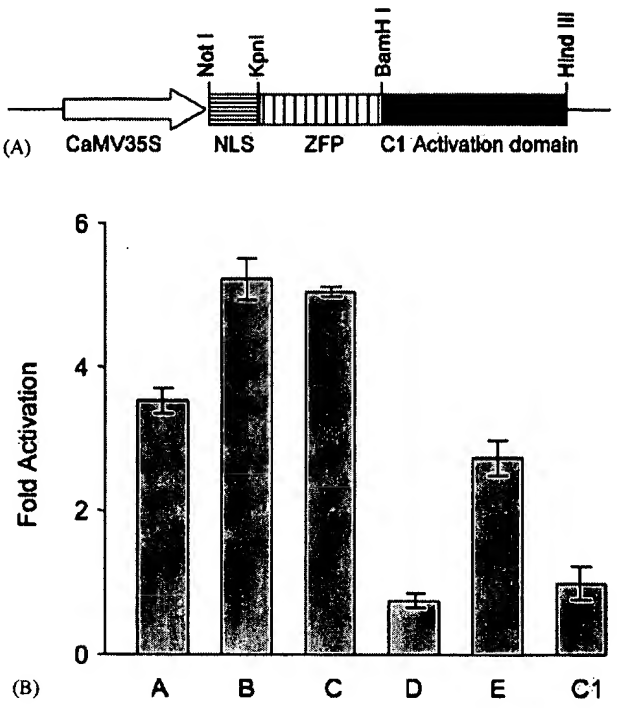


Fig. 4. Transcriptional activation of the endogenous *Arabidopsis* GMT gene by ZFP-TFs in leaf protoplasts. (A) Schematic representation of the protoplast ZFP-TF expression cassette. The cauliflower mosaic virus 35S (CaMV35S) constitutive promoter was used to drive the expression of the ZFP-TF fusion protein (maize opaque-2 nuclear localization signal (NLS), plant-derived ZFP DNA binding domain, and maize C1 activation domain). (B) Differential activation of the GMT gene by ZFP-TFs (A–E) targeted to different sites on the GMT locus. Plasmids encoding ZFP-TFs were transformed into *Arabidopsis* leaf protoplasts. Taqman quantitative RT-PCR analysis was performed on RNA extracted from protoplasts. GMT mRNA level was normalized against GAPDH mRNA and the transformation efficiency. The normalized GMT mRNA level was compared to the control vector containing only the CaMV35S driven C1 activation domain.

3.4. Phenotypic analysis of transgenic plants

The five ZFP-TFs were put under the control of the napin embryo-specific promoter and stably transformed into *Arabidopsis*. Tocopherol compositional analysis via HPLC from transgenic lines showed that segregating T2 seed contained as much as 19.8% α-tocopherol compared to an average of 1% α-tocopherol in control seed (Table 1) while total tocopherol content was unchanged (data not shown). The relative in planta efficacy of the ZFP-TFs to increase seed α-tocopherol percentage was B > A > E > D > C. There was a lot of variability in the seed α-tocopherol percentage from different lines expressing the same ZFP-TFs. Analysis of T3 seed from transgenic lines expressing either ZFP-TF A or B showed that the elevated α-tocopherol phenotype was heritable, with 26.9% being the highest α-tocopherol level recorded in T3 seed (Fig. 5). Although levels of α-

tocopherol were significantly elevated in several lines, the amount was less than that seen when the GMT cDNA itself was expressed under the control of a seed-specific promoter (Shintani et al., 1998). Significantly, the two ZFP-TFs that were targeted to DNase I hypersensitive sites (A and B) gave the highest increase in seed α-tocopherol percentage. C and E were effective activators in leaf but not in seed suggesting that these sites may not be accessible in the embryo. Hypersensitivity mapping is best performed in the target tissue. We chose to use leaf tissue in this study as we determined that it was impractical to collect a sufficient quantity of Arabidopsis embryo target tissue. Given that the GMT gene is normally poorly expressed in seed relative to the leaf it may therefore have a less accessible chromatin configuration in seed. Although it is apparent that the effectiveness of a ZFP-TF in one tissue does not guarantee its effectiveness in another, the finding that the two ZFP-TFs targeted to the major hypersensitive sites were effective in both leaf and seed tissue suggests that it is feasible to use DNaseI hypersensitivity mapping of a given gene in one tissue to predict accessible sites of that gene in another more scarce tissue type.

Table 1
Alpha-tocopherol percentage in mature T2 seed and frequency of zinc finger protein transcription factor (ZFP-TF) transgene expression in developing siliques from T1 transgenic *Arabidopsis* plants

Construct	Seed α-tocopherol (%)					Transgene expression	
	N	Mean	SEM	Min	Max	No. expressing/ no. assayed	%
Control[a]	30	1.0	0.03	0.5	1.5	0/17	0
A	18	4.5	0.68	1.3	8.8	7/13	54
B	13	6.0	1.59	1.0	19.8	4/7	57
C	35	1.6	0.03	1.1	2.1	16/16	100
D	12	2.3	0.12	1.8	3.2	12/12	100
E	22	2.9	0.23	1.3	5.5	14/17	82

[a] Controls were transformed with the parent binary vector (lacking a ZFP-TF insert).

3.5. Expression analysis of transgenic plants

TaqMan quantitative RT-PCR analysis on developing siliques containing segregating T2 seed showed good concordance between the presence of the ZFP-TF transgene and an elevated level of seed α-tocopherol, however not all transgene expressors had elevated seed α-tocopherol (Table 1). The endogenous napin mRNA level was high in all samples, confirming that the developmental stage being assayed corresponded to the time when the napin-driven transgene was expected to be expressed (data not shown). Twenty plants transformed with the two most effective ZFP-TFs in planta (A and B) and nine plants transformed with a control vector (binary transformation vector lacking a ZFP-TF insert) were selected for detailed expression analysis of C1-ED and endogenous GMT levels. Within these 20 samples there was concordance between the presence of transgene expression and elevated α-tocopherol percentage in the seed, but no correlation between relative expression level of the ZFP-TFs transgene in developing siliques and percentage α-tocopherol in mature seed of the expressors (Fig. 6). This lack of correlation may be explained by the observation that only a few copies of the ZFP-TFs would be required to completely saturate

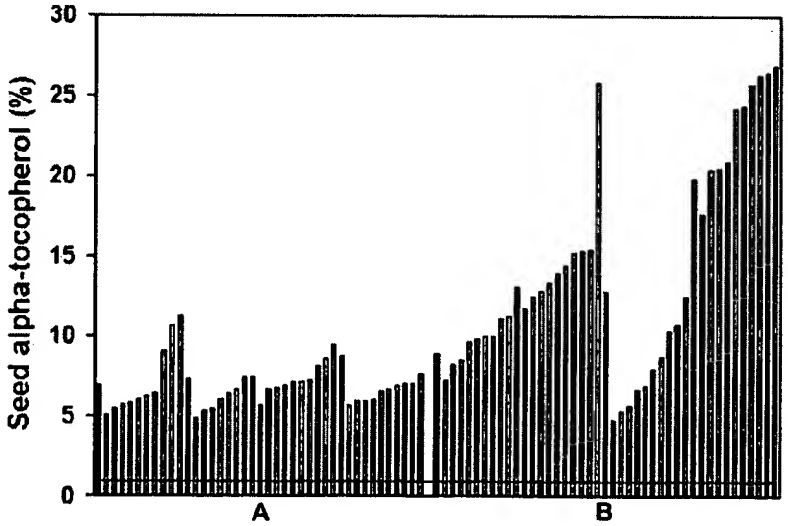


Fig. 5. α-tocopherol percentage in T2 and T3 seed from selected transgenic lines. α-tocopherol percentage in seed from individual T1 (black bar) and several T2 (gray bars) *Arabidopsis* plants derived from four independent transformation events with one of two (A or B) ZFP-TFs targeted to different recognition sites on the endogenous GMT gene. Control value of 1.0 (black horizontal line; SEM 0.03) represents the average α-tocopherol percentage measured in control seed samples from 30 plants transformed with the parent binary vector (lacking a ZFP-TF insert).

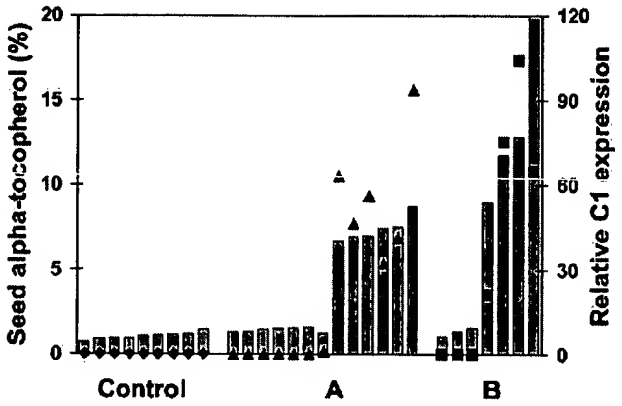


Fig. 6. α-tocopherol percentage and relative expression of ZFP-TF transgenes in developing siliques from T1 transgenic *Arabidopsis* plants. α-tocopherol percentage in T2 seed (bars) and relative expression levels of transgenic ZFP-TFs in developing siliques containing T2 segregating seed from individual transgenic plants transformed with a binary vector containing no ZFP-TF (diamond), ZFP-TF A (triangle), or ZFP-TF B (square) under the control of the napin embryo-specific promoter. Transgene expression data are normalized with 18S RNA and displayed as fold induction relative to the expression level of the weakest C1 expressor (arbitrarily set to a value of 1).

the specific binding sites, based on the size of the nucleus and the calculated in vitro dissociation constants of the ZFPs (Fig. 4). This calculation is complicated by the fact that the ZFPs may recognize other non-target sites in the nucleus. This would differentially increase the minimally effective concentration for GMT activation for each of the ZFP-TFs examined here.

There was no detectable increase in the quantity of endogenous GMT transcripts in developing siliques from these transgenic plants (data not shown). One likely explanation is that the TaqMan analysis utilized whole siliques due to the difficulty of isolating embryos from developing siliques. GMT is expressed in green, photosynthetic tissue and the high background of GMT expression in the silique walls may have made it difficult to pick up the transcriptional activation of GMT expression in transgenic embryos. It may also be that napin promoter expression of the ZFP-TFs caused an alteration in the temporal presence of endogenous GMT transcript that was not apparent in our analysis. Further expression analysis on isolated embryos from different time points in seed development would be required to assess the timing and magnitude of the transcriptional effect.

The 9 bp ZFP target sequences we chose were not expected to be unique within the *Arabidopsis* genome, and would be predicted to occur an average of every 2.6×10^5 bp. A search of the 1000 bp region upstream of all of the predicted *Arabidopsis* ORFs at The *Arabidopsis* Information Resource (TAIR) website (Huala et al., 2001) (www.arabidopsis.org/tools/) found the binding sequence of A upstream of 127 ORFs, B upstream of 96 ORFs and E upstream of 694 ORFs. Although chromatin infrastructure would likely prevent the binding of the ZFP-TFs to most of these sites, it is unlikely that all sites would be inaccessible. We had some difficulty obtaining expressing transgenic lines in three (A, B and E) of the five ZFP-TFs examined suggesting pleiotropic effects were occurring with these constructs. Using ZFP-TFs with more fingers targeting a longer DNA binding site would be one way to ensure the specificity of activation (Tan et al., 2003) and ameliorate some of these apparent pleiotropic effects.

4. Conclusion

In this study we were able to elevate the vitamin E content of transgenic *Arabidopsis* seed by expressing ZFP-TFs designed to activate the endogenous GMT gene. Our experimental approach incorporated information regarding the chromatin structure of the endogenous GMT locus and utilized plant-derived DNA sequences for the design of synthetic ZFP-TFs. Five ZFPs were designed to 9 bp DNA sequences in the promoter or coding region of the GMT gene and all were found to bind with strong affinity to their naked DNA target. ZFPs were fused to the maize C1 activation domain and four of these synthetic ZFP-TFs were found to upregulate the expression of the endogenous GMT gene in a leaf protoplast assay system. It was further shown that these ZFP-TFs were able to alter seed tocopherol composition when expressed in transgenic plants. The best line demonstrated a dramatic and heritable 20+ fold increase in seed α-tocopherol percentage. Significantly, the two ZFP-TFs that were targeted to the major DNase I hypersensitive sites in the GMT promoter gave the highest increase in seed α-tocopherol percentage. These results demonstrate that designer ZFP-TFs can be used to target endogenous gene expression and increase the amount of product derived from a specific step in a native biochemical pathway. ZFP-TF technology provides agricultural biotechnologists a powerful new method for modifying endogenous plant gene expression to achieve desired phenotypic alterations.

Acknowledgments

Thanks to Alan Wolffe for his guidance in the initial phases of this project, Michelle Kong for DNase I hypersensitive mapping, Nancy Foster for ZFP assembly, Katherine Howes for gel shift analysis of ZFPs, Steve Strickland and Brenda Reed for *Arabidopsis* plant transformation and maintenance, Susan Baszis for tocopherol analysis of *Arabidopsis* seed, and JoAnne

Fillatti for helpful input on the project and manuscript review.

References

Abel, S., Theologis, A., 1994. Transient transformation of *Arabidopsis* leaf protoplasts: a versatile experimental system to study gene expression. Plant J. 5, 421–427.

Bartsevich, V.V., Juliano, R.L., 2000. Regulation of the MDR1 gene by transcriptional repressors selected using peptide combinatorial libraries. Mol. Pharmacol. 58, 1–10.

Beerli, R.B., Barbas III, C.F., 2002. Engineering polydactyl zinc-finger transcription factors. Nat. Biotechnol. 20, 135–141.

Beerli, R.R., Drier, B., Barbas III, C.F., 2000. Positive and negative regulation of endogenous genes by designed transcription factors. Proc. Natl. Acad. Sci. USA 97, 1495–1500.

Bramley, P.M., Elmadfa, I., Kafatos, A., Kelly, F.J., Manios, Y., Roxborough, H.E., Schuch, W., Sheehy, P.J.A., Wagner, K.-H., 2000. Vitamin E. J. Sci. Food. Agric. 80, 913–938.

Clough, S., Bent, A., 1998. Floral dip: a simplified method for *agrobacterium*-mediated transformation of *Arabidopsis thaliana*. Plant J. 16, 735–743.

Cook, T., Gebelein, B., Urrutia, R., 1999. Sp1 and its likes: biochemical and functional predictions for a growing family of zinc finger transcription factors. Ann. N. Y. Acad. Sci. 880, 94–102.

Dreier, B., Segal, D.J., Barbas III, C.F., 2001. Development of zinc finger domains for recognition of the 5′-ANN-3′ family of dna sequences and their use in the construction of artificial transcription factors. J. Biol. Chem. 276, 29466–29478.

Goff, S.A., Cone, K.C., Fromm, M.E., 1991. Identification of functional domains in the maize transcriptional activator C1: comparison of wild-type and dominant inhibitor proteins. Genes Dev. 5, 298–309.

Guan, X., Stege, J., Kim, M., Dahmani, Z., Fan, N., Heifetz, P., Barbas III, C.F., Briggs, S.P., 2002. Heritable endogenous gene regulation in plants with designed polydactyl zinc finger transcription factors. Proc. Natl. Acad. Sci. USA 99, 13296–13301.

Guyer, D., Tuttle, A., Rouse, S., Volrath, S., Johnson, M., Potter, S., Görlach, J., Goff, S., Crossland, L., Ward, E., 1998. Activation of latent transgenes in *Arabidopsis* using a hybrid transcription factor. Genetics 149, 633–639.

Huala, E., Dickerman, A.W., Garcia-Hernandez, M., Weems, D., Reiser, S., LaFond, F., Hanley, D., Kiphart, D., Zhuang, M., Huang, W., Mueller, L.A., Bhattacharyya, D., Bhaya, D., Sobral, B.W., Beavis, W., Meinke, D.W., Town, C.D., Somerville, C., Rhee, S.Y., 2001. The *Arabidopsis* information resource (TAIR): a comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. Nucleic Acids Res. 29, 102–105.

Jamieson, A., Li, G., 2002. Zinc finger proteins for DNA binding and gene regulation in plants. International Patent application no. WO 02/057294.

Kridl, J., McCarter, D.W., Rose, R.E., Scherer, D.E., Knutzon, D.S., Radke, S.E., Knauf, V.C., 1991. Isolation and characterization of an expressed napin gene from *Brassica napus*. Seed Sci. Res. 1, 209–219.

Li, G., Chandler, S., Wolffe, A., Hall, T., 1998. Architectural specificity in chromatin structure at the TATA box in vivo: nucleosome displacement upon beta-phaseolin gene activation. Proc. Natl. Acad. Sci. USA 95, 4772–4777.

Li, G., Chandrasekharan, M., Wolffe, A., Hall, T., 2001. Chromatin structure and phaseolin gene regulation. Plant Mol. Biol. 46, 121–129.

Liu, P.-Q., Rebar, E.J., Zhang, L., Liu, Q., Jamieson, A.C., Liang, Y., Qi, H., Li, P.-X., Chen, B., Mendel, M.C., Zhong, X., Lee, Y.-L., Eisenberg, S.P., Spratt, S.K., Case, C.C., Wolffe, A.P., 2001. Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions. J. Biol. Chem. 276, 11323–11334.

McBride, K.E., Schaaf, D.J., Daley, M., Stalker, D.M., 1994. Controlled expression of plastid transgenes in plants based on a nuclear DNA-encoded and plastid-targeted T7 RNA polymerase. Proc. Natl. Acad. Sci. USA 91, 7301–7305.

Ordiz, M.I., Barbas III, C.F., Beachy, R.N., 2002. Regulation of transgene expression in plants with polydactyl zinc finger transcription factors. Proc. Natl. Acad. Sci. USA 99, 13290–13295.

Pabo, C., Peisach, E., Grant, R., 2001. Design and selection of novel Cys_2His_2 zinc finger proteins. Annu. Rev. Biochem. 70, 313–340.

The Perkin–Elner Corporation (PE_Applied_Biosytems), 1997. User Bulletin #2. Relative quantitation of gene expression. ABI prism 7700 sequence detections system. The Perkin–Elmer Corporation, Chicago, IL.

Rebar, E., Jamieson, A., 2002. Modified zinc finger binding proteins. International patent application publication no. WO 02/057293.

Ren, D., Collingwood, T.N., Rebar, E.J., Wolffe, A.P., Camp, H.S., 2002. PPARgamma knockdown by engineered transcription factors: exogenous PPARgamma 2 but not PPARgamma 1 reactivates adipogenesis. Genes Dev. 16, 27–32.

Savidge, B., Weiss, J.D., Wong, Y.-H.H, Lassner, M.W., Mitsky, T.A., Shewmaker, C.K., Post-Beittenmiller, D., Valentin, H.E., 2002. Isolation and characterization of homogentisate phytyltransferase genes from *Synechocystis* PCC 6803 and *Arabidopsis*. Plant Physiol. 129, 321–332.

Segal, D.J., Barbas III, C.F., 2001. Custom DNA-binding proteins come of age: polydactyl zinc-finger proteins. Curr. Opin. Biotechnol. 12, 632–637.

Sheppard, A.J., Pennington, J.A., Weihrauch, J.L., 1993. In: Packer, J.F.L. (Ed.), Vitamin E in health and disease. Marcel Dekker, New York, pp. 9–31.

Shintani, D., Della-Penna, D., 1998. Elevating the vitamin E content of plants through metabolic engineering. Science 282, 2098–2100.

Tan, S., Gustin, D., Davalos, A., Lee, Y.-L., Snowden, A.W., Jouvenot, Y., Zhang, H.S., Howes, K., McNamara, A.R., Lai, A., Ullman, C., Reynolds, L., Moore, M., Isalan, M., Berg, L.-P., Campos, B., Qi, H., Spratt, K., Case, C.C., Pabo, C.O., Campisi, J., Gregory, P.D., 2003. Zinc-finger protein targeted gene regulation: genomewide single-gene specificity. Proc. Natl. Acad. Sci. USA 100, 11997–12002.

Utley, R.T., Cote, J., Owen-Hughes, T., Workman, J.L., 1997. SWI/SNF stimulates the formation of disparate activator-nucleosome complexes but is partially redundant with cooperative binding. J. Biol. Chem. 272, 12642–12649.

Zhang, L., Spratt, S.K., Liu, Q., Johnstone, B., Qi, H., Raschke, E.E., Jamieson, A.C., Rebar, E.J., Wolffe, A.P., Case, C.C., 2000. Synthetic zinc finger transcription factor action at an endogenous chromosomal site. J. Biol. Chem. 275, 33850–33860.